A CONNECTION BETWEEN MIXING AND KAC'S CHAOS

GEORGE ANDROULAKIS AND RADE MUSULIN

Abstract. The Boltzmann equation is an integro-differential equation which describes the density function of the distribution of the velocities of the molecules of dilute monoatomic gases under the assumption that the energy is only transferred via collisions between the molecules. In 1956 Kac studied the Boltzmann equation and defined a property of the density function that he called the "Boltzmann property" which describes the behavior of the density function at a given fixed time as the number of particles tends to infinity. The Boltzmann property has been studied extensively since then, and now it is simply called chaos, or Kac's chaos. On the other hand, in ergodic theory, chaos usually refers to the mixing properties of a dynamical system as time tends to infinity. A relationship is derived between Kac's chaos and the notion of mixing.

1. Several notions of chaos

The notion of "chaos" in ergodic theory, has its origins in the works of Poincare at the end of the 19th century. Its meaning is dynamical randomness of physical quantities that evolve with time. The set up for measure theoretic dynamical systems consists of a probability space (Ω, Σ, μ) which is called the *phase space*, and either a measurable map $S : \Omega \to \Omega$ (in the case of discrete time dynamical systems), or a family of measurable maps $S_t : \Omega \to \Omega$ for $t \geq 0$ (in the case of continuous time dynamical systems) satisfying $S_t \circ S_s = S_{t+s}$ for all $s, t \in [0, \infty)$ (semigroup property). Such tuple (Ω, Σ, μ, S) or $(\Omega, \Sigma, \mu, (S_t)_{t\geq 0})$ is called a *measure theoretic dynamical system*, or simply a *dynamical system*. In the case of discrete time dynamical systems, the composition of the map S with itself n many times, (where n is a non-negative integer), is usually denoted as S^n, (a notation which resembles powers of S), and plays the role of S_n that appears in the above definition of continuous time dynamical systems. For simplicity we only consider discrete time dynamical systems and we keep in mind that the "exponent" n that appears in the compositions S^n represents time. The maps S^n can be thought to act on Ω, (by the formula $\Omega \ni \omega \mapsto S^n\omega$), or on real valued functions on Ω, (where the action of S^n on such function f produces the real valued function

1991 *Mathematics Subject Classification.* Primary: 37A25, Secondary: 81Q50, 28D05.
Key words and phrases. mixing, chaos, Kac's chaos.
The article is part of the second author's Ph.D. thesis which is prepared at the University of South Carolina under the supervision of the first author.

$\Omega \ni \omega \mapsto f(S^n\omega))$, or on probability measures on Σ (where the action of S^n on such measure ν produces the measure $\Sigma \ni A \mapsto \nu(S^{-n}A))$. Thus one can study orbits of points of Ω, (i.e. the sequence of points $(S^n(\omega))_{n\in\mathbb{N}\cup\{0\}})$, or orbits of real valued functions on Ω, or orbits of probability measures on Σ. The property of chaos in ergodic theory refers to the randomness of these orbits and it is explicitly quantified and studied in the books of ergodic theory. An excellent book on this subject is the book of Arnold and Avez, [1], or the short survey of Sinai [21]. Two quantifications of the notions of chaos in the measure theoretic ergodic theory are the notions of the "stationary limit" and "mixing":

Definition 1.1. (i) *We say that a dynamical system (Ω, Σ, μ, S) is asymptotically stationary with stationary limit ν if $\nu(A) = \lim_{k\to\infty} \mu(S^{-k}A)$ for each $A \in \Sigma$.*

(ii) *We say that a dynamical system (Ω, Σ, μ, S) is mixing if*

$$\lim_{k\to\infty} \left|\mu(S^{-k}A\cap B) - \mu(S^{-k}A)\mu(B)\right| = 0 \quad \textit{for all } A, B \in \Sigma.$$

Note that if the members of a sequence of probability measures are defined on a common σ-algebra Σ and converge at every fixed element of Σ then the limit is also a probability measure [2, Theorem 4.6.3(i)]. Thus the limit ν that is obtained in Definition 1.1(i) is a probability measure, since obviously, for every $k \in \mathbb{N}$, the map $\Sigma \ni A \mapsto \mu(S^{-k}A)$ defines a probability measure on Σ. Obviously, if a dynamical system (Ω, Σ, μ, S) is asymptotically stationary with stationary limit ν then ν is *invariant under* S, (or equivalently, S *is* ν-*measure preserving*), i.e.

$$\nu(S^{-1}(A)) = \nu(A) \quad \text{for all } A \in \Sigma. \tag{1}$$

It is also obvious that if the dynamical system (Ω, Σ, μ, S) is asymptotically stationary with stationary limit ν then it is mixing if and only if

$$\lim_{k\to\infty} \left|\mu(S^{-k}A\cap B) - \nu(A)\mu(B)\right| = 0 \quad \text{for all } A, B \in \Sigma.$$

In particular, if (Ω, Σ, μ, S) is a dynamical system and the map S is μ-measure preserving, then (Ω, Σ, μ, S) is mixing if and only if

$$\lim_{k\to\infty} \left|\mu(S^{-k}A\cap B) - \mu(A)\mu(B)\right| = 0 \quad \text{for all } A, B \in \Sigma.$$

Another notion of chaos was created in 1956 by Kac [13] while he was studying the Boltzmann equation. For a fixed positive integer n, the Boltzmann equation describes the density function of the distribution of the velocities of n many molecules of dilute monoatomic gases where the energy is assumed to be transferred only via elastic collisions between the molecules. While the Boltzmann equation is a non-linear equation, Kac came up with a linear integro-differential equation that he called the "master equation" [13, Equation (2.6)]. Kac's master equation depends on a positive integer n and its solution has $n + 1$ real variables $(x_1, \ldots, x_n, t)$. The n real variables $(x_1, \ldots, x_n)$ belong on the "Kac's sphere" $\mathbb{K}^n$ which stands for the sphere in $\mathbb{R}^n$ centered at the origin whose radius is equal to $\sqrt{n}$, (i.e. $\mathbb{K}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = n\}$), while the extra variable t represents time. Kac seeked solutions $f^{(n)}$ to his master equation that are symmetric in the variables $(x_1, \ldots, x_n)$ for every $t \geq 0$, i.e.

$$f^{(n)}(x_1 \ldots, x_n, t) = f^{(n)}(x_{\pi(1)}, \ldots, x_{\pi(n)}, t), \quad \text{for every permutation } \pi \text{ of } \{1, \ldots, n\}. \tag{2}$$

Moreover for any set E, a function $g : E^n \to \mathbb{C}$ is called symmetric if

$$g^\pi(x_1, x_2, ..., x_n) = g(x_1, x_2, ..., x_n) \tag{3}$$

for all permutations π of $\{1, \ldots, n\}$ and for all $(x_1, x_2, ..., x_n) \in E^n$, where for each permutation π of $\{1, \ldots, n\}$ we define $g^\pi : E^n \to \mathbb{C}$ by

$$g^\pi(x_1, x_2, ..., x_n) := g(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n)}). \tag{4}$$

We also assume that the solution $f^{(n)}$ to Kac's master equation is a *density function on* $\mathbb{K}^n$ i.e. we assume that $f^{(n)} d\sigma^n$ is a probability measure on the Borel subsets of $\mathbb{K}^n$ where σ^n denotes the normalized uniform measure on $\mathbb{K}^n$. For each $1 \leq m \leq n$, we can define a probability measure $(f^{(n)} d\sigma^n)_m$ on the Borel subsets of $\mathbb{R}^m$ by

$$(f^{(n)} d\sigma^n)_m(A) = \int_{P_m^{-1}(A)} f^{(n)} d\sigma^n$$

where $P_m : \mathbb{K}^n \to \mathbb{R}^m$ is the canonical projection into the first m copies of $\mathbb{R}$. It is clear that $(f^{(n)} d\sigma^n)_m$ is absolutely continuous with respect to the m-dimensional Lebesgue measure λ^m on $\mathbb{R}^m$, and thus by the Radon-Nikodym Theorem there exists a function $f_m^{(n)} \in L^1(\mathbb{R}^m)$, called the m^{th} marginal function of $f^{(n)}$, such that

$$\int_A f_m^{(n)} d\lambda^m = (f^{(n)} d\sigma^n)_m(A) = \int_{P_m^{-1}(A)} f^{(n)} d\sigma^n$$

for every Borel subset A of $\mathbb{R}^m$. Hence $f_m^{(n)} d\lambda^m$ is a Borel probability measure on $\mathbb{R}^m$ for every $1 \leq m \leq n$. Kac observed that if $f^{(n)}$ satisfies the master equation [13, Equation (2.6)] then the first and second marginals $f_1^{(n)}$ and $f_2^{(n)}$ satisfy [13, Equation (3.7)] which reads

$$\frac{\partial f_1^{(n)}(x,t)}{\partial t} = \frac{(n-1)\nu}{2\pi n} \int_{-\sqrt{n-x^2}}^{\sqrt{n-x^2}} f_2^{(n)}(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, t) - f_2^{(n)}(x, y, t) d\theta dy. \tag{5}$$

This equality is interpreted in the weak sense for density functions, i.e. each side is integrated against smooth functions with compact support having variables $(x, t) \in \mathbb{R} \times [0, \infty)$. In particular, the derivative is interpreted in the sense of distributions.

Kac observed that if the limits $f_1(\cdot, t) := \lim_{n\to\infty} f_1^{(n)}(\cdot, t)$ and $f_2(\cdot, \cdot, t) := \lim_{n\to\infty} f_2^{(n)}(\cdot, \cdot, t)$ exist in $L^1(\mathbb{R})$ and $L^1(\mathbb{R}^2)$ respectively, for all $t \geq 0$, (where the dots denote arbitrary real variables, and the L^1 spaces are taken with respect to the Lebesgue measure), and if for almost all $x, y \in \mathbb{R}$,

$$f_2(x, y, t) = f_1(x, t) f_1(y, t), \tag{6}$$

then

$$\frac{\partial f_1(x,t)}{\partial t} = \frac{\nu}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} f_1(x\cos\theta + y\sin\theta, t) f_1(-x\sin\theta + y\cos\theta, t) - f_1(x,t) f_1(y,t) d\theta dy \tag{7}$$

again, in the weak sense, i.e. the function f_1 is a solution to a simplified version of the non-linear Boltzmann equation.

Equation (6) motivated Kac to introduce the following definition: For all $n \in \mathbb{N}$ let $f^{(n)}$ be a symmetric, (as in (2)), probability density function defined on $\mathbb{K}^n$ (i.e. $f^{(n)} d\sigma^n$ is a Borel probability measure on $\mathbb{K}^n$). For $1 \leq k \leq n$ let $f_k^{(n)}$ denote the k^{th} marginal of $f^{(n)}$. The sequence $(f^{(n)})$ is said to have the "Boltzmann property" if for all $k \in \mathbb{N}$ the limit $\lim_{n\to\infty} f_k^{(n)}$ exists in $L^1(\mathbb{R}^k)$, and moreover, if f_1 denotes the $L^1(\mathbb{R})$ limit of $f_1^{(n)}$, then for all $k \in \mathbb{N}$ and for almost all $x_1, \ldots, x_k \in \mathbb{R}$:

$$\lim_{n\to\infty} f_k^{(n)}(x_1, \ldots, x_k) = \prod_{i=1}^{k} f_1(x_i). \tag{8}$$

Kac proved that if the initial value solution (at time $t = 0$) to the master equation is symmetric (as in (2)) and satisfies the Boltzmann property then the solution to the master equation is symmetric and satisfies the Boltzmann property for all times $t > 0$. Since Kac's master equation is linear hence the existence of its solution is guaranteed by well known theory, Kac produced a method for constructing a solution to a simplified version of the non-linear Boltzmann equation.

The "Boltzmann property" is commonly referred to as "Kac's chaos" and has attracted the interest of many people such as McKean [17], Johnson [12], Tanaka [24], Ueno [25], Grünbaum [9], Murata [20], Graham and Méléard [8], Sznitman [22], [23], Mischler [18], Carlen, Carvalho and Loss [4], Michler and Mouhot [19]. These authors considered a more general situation than a sequence $f^{(n)}$ of density functions defined on $\mathbb{K}^n$. They considered a topological space E and a symmetric probability measure μ_n on the Borel σ-algebra $\mathcal{B}(E^n)$ of the Cartesian product E^n for each $n \in \mathbb{N}$. Here are the relevant definitions, where for a topological space E, we denote by $C_b(E)$ the set of all continuous bounded functions on E:

Definition 1.2. *Let E be a topological space, n be a positive integer, μ_n be a probability measure on the Borel subsets of E^n. Then μ_n is called symmetric if for any $\phi_1, \phi_2, ..., \phi_n \in$*

$C_b(E)$,

$$\int_{E^n} \phi_1(x_1)\phi_2(x_2)\cdots\phi_n(x_n)d\mu_n = \int_{E^n} \phi_1(x_{\pi(1)})\phi_2(x_{\pi(2)})\cdots\phi_n(x_{\pi(n)})d\mu_n$$

for any permutation π of $\{1,\ldots,n\}$.

Note that if $f^{(n)}$ is a symmetric (in the sense of (2)) density function on the Kac's sphere $\mathbb{K}^n$, σ^n denotes, as above, the uniform Borel probability measure on $\mathbb{K}^n$, and $\widetilde{\sigma^n}$ denotes the extension of σ^n to the Borel subsets of $\mathbb{R}^n$ such that the support of $\widetilde{\sigma^n}$ is equal to $\mathbb{K}^n$ (this is possible since $\mathbb{K}^n$ is a Borel subset of $\mathbb{R}^n$), then $f^{(n)}d\widetilde{\sigma^n}$ is a symmetric probability measure on $\mathbb{R}^n$, (in the sense of Definition 1.2). Thus Definition 1.2 gives a more general notion of symmetry than that of Equation (2) that was considered by Kac. Now we define the Boltzmann property, or Kac's chaos, but following the above mentioned literature, we use a more descriptive terminology:

Definition 1.3. *Let E be a topological space, ν be a Borel probability measure on E, and for every $n \in \mathbb{N}$ let μ_n be a symmetric (as in Definition 1.2) Borel probability measure on E^n. We say that $(\mu_n)_{n=1}^{\infty}$ is ν-chaotic if for all $k \geq 1$ and $\phi_1, \phi_2, ..., \phi_k \in C_b(E)$,*

$$\lim_{n\to\infty} \int_{E^n} \phi_1(x_1)\phi_2(x_2)\cdots\phi_k(x_k)d\mu_n = \prod_{j=1}^{k} \int_E \phi_j(x)d\nu(x).$$

Now let $f^{(n)}$ be a symmetric (in the sense of (2)) density on $\mathbb{K}^n$ for all $n \in \mathbb{N}$ such that the sequence $(f^{(n)})_n$ has the Boltzmann property (as defined by Kac). In particular, let f_1 be the $L^1(\mathbb{R})$ limit of the sequence $(f_1^{(n)})_n$. Extend each $f^{(n)}$ to $\mathbb{R}^n$ (without changing its name) by setting it equal to zero on $\mathbb{R}^n \backslash \mathbb{K}^n$, and let $\widetilde{\sigma^n}$ be the Borel probability measure on $\mathbb{R}^n$ which is supported on $\mathbb{K}^n$ and it is uniform on $\mathbb{K}^n$. Then the sequence of measures $(f^{(n)}d\widetilde{\sigma^n})_n$ defined on the Borel subsets of $\mathbb{R}^n$ is ν-chaotic where $d\nu = f_1dx$, and dx is the Lebesgue measure on $\mathbb{R}$. Thus Definition 1.3 gives a more general notion of chaoticity than the Boltzmann property defined by Kac.

In this article we provide a relationship between the notion of mixing that appears in ergodic theory and Kac's chaos. Our main result is Theorem 2.1 which asserts that given a dynamical system on a separable metric space E which satisfies a property similar to the mixing property and it is asymptotically stationary with stationary limit ν, one can construct a sequence of Borel probability measures $(\mu_n)_n$ on $(E^n)_n$ which is ν-chaotic.

Two other related forms of chaoticity that exist in literature are the chaoticity in the sense of Boltzmann entropy and the chaoticity in the sense of Fisher information. These two notions were introduced by Carlen, Carvalho, Le Roux, Loss, and Villani [3]. Hauray and Mischler has shown that chaoticity in the sense of Fisher information implies chaoticity in the sense of Boltzmann entropy, which in turn implies Kac's chaoticity [10, Theorem 1.4].

Carrapatoso [5] has extended the results of [10] to probability measures with support on the Boltzmann spheres.

Finally we would like to mention that there is a vast literature on the notion of "quantum chaos", where notions of ergodic theory are extended to quantum physical models. Without attempting to give detailed references to quantum chaos, we refer the interested reader to the books [6], [7], and [11] where some of these notions are presented.

2. The statement of the main result and some examples

In this section we state the main result of the article and we give several examples of dynamical systems that satisfy its assumptions. Before stating the main result we introduce some notation. If E is a topological space then $\mathcal{B}(E)$ will denote the σ-algebra of the Borel subsets of E, and $M(E)$ will denote the set of probability measures on $\mathcal{B}(E)$. Also Σ_n will denote the set of permutations of $\{1,\dots,n\}$ for each $n \in \mathbb{N}$.

We now present the main result of the article.

Theorem 2.1. *Let E be a separable metric space, μ be a probability measure on $\mathcal{B}(E)$, and $S : E \to E$ be a Borel measurable map. Assume that*

1. *For every $A \in \mathcal{B}(E)$,*
$$\sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}A) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}A)| \xrightarrow[k\to\infty]{} 0,$$
and
2. *$(E, \mathcal{B}(E), \mu, S)$ is asymptotically stationary with stationary limit ν.*

For every $n \in \mathbb{N}$ define $\mu_n : \mathcal{B}(E^n) \to [0,1]$ by
$$\mu_n(A) = \frac{1}{n!} \sum_{\sigma\in\Sigma_n} \mu\{x \in E : (S^{\sigma(1)}(x), ..., S^{\sigma(n)}(x)) \in A\}.$$
Then $(\mu_n)_{n\in\mathbb{N}}$ is ν-chaotic.

Note that assumption 1 of the main result is related to the mixing property that was introduced in Definition 1.1(ii). The differences between the two properties are: The limit in Definition 1.1(ii) is taken for any Borel sets A and B while in assumption 1, the sets A and B are equal and they belong to the σ-algebra $\{S^{-i}A : A \in \mathcal{B}(E)\}$. In that sense, assumption 1 is weaker than Definition 1.1(ii). On the other hand, Definition 1.1(ii) lacks the uniformity which is manifested in assumption 1 by the presence of the supremum. In that sense, assumption 1 is stronger than Definition 1.1(ii).

One way to guarantee that a dynamical system satisfies assumption 1 of Theorem 2.1 is by means of the next lemma.

Lemma 2.2. *Let (Ω, Σ, μ, S) be a dynamical system which is asymptotically stationary. Let $\Pi \subset \Sigma$ be a π-system such that $\sigma(\Pi) = \Sigma$ and*

$$\lim_{k\to\infty} \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| = 0 \tag{9}$$

for all $A, B \in \Pi$. Then (9) is satisfied for all $A, B \in \Sigma$ (hence assumption 1 of Theorem 2.1 is satisfied as well).

Proof. The proof will be by way of the Dynkin $\pi - \lambda$ Theorem. Fix $A \in \Pi$, and define $\Lambda_A := \{B \in \Sigma : \lim_{k\to\infty} \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| = 0\}$. It is clear that $\Omega \in \Lambda_A$. Now, assume $B_1, B_2 \in \Lambda_A$ such that $B_1 \subset B_2$. Then we have

$$\begin{aligned}
0 &\leq \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}(B_2 \setminus B_1)) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}(B_2 \setminus B_1))| \\
&= \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_2) - \mu(S^{-i}A \cap S^{-k}S^{-i}B_1) \\
&- \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_2) + \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_1)| \\
&\leq \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_2) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_2)| \\
&+ \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_1) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_1)|
\end{aligned}$$

Taking limits on both sides as $k \to \infty$, we get that $B_2 \setminus B_1 \in \Lambda_A$.

Lastly, let $(B_n)_{n=1}^{\infty} \subset \Lambda_A$ be any monotone increasing sequence with limit $B \in \Sigma$. We need to show that $B \in \Lambda_A$. Let $\epsilon > 0$. Denote by ν the stationary limit of (Ω, Σ, μ, S). We have that

$$\begin{aligned}
&\sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| \\
&\leq \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\nu(S^{-i}B)| + \sup_{i\in\mathbb{N}} |\mu(S^{-i}A)||\nu(B) - \mu(S^{-k}S^{-i}B)|
\end{aligned} \tag{10}$$

By assumption, there exists a $k_0 \in \mathbb{N}$ such that $|\nu(B) - \mu(S^{-k}B)| < \epsilon$ for all $k \geq k_0$, and thus, the second term of (10) can be made small. Notice that there exists an $n_0 \in \mathbb{N}$ such that $|\nu(B_n) - \nu(B)| < \epsilon$ for all $n \geq n_0$. Using this information, we focus on the first term of line (10),

$$
\begin{aligned}
& \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\nu(S^{-i}B)| \\
= \; & \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_{n_0}) + \mu(S^{-i}A \cap S^{-k}S^{-i}(B \setminus B_{n_0})) - \mu(S^{-i}A)\nu(S^{-i}B)| \\
\le \; & \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_{n_0}) - \mu(S^{-i}A)\nu(S^{-i}B)| + \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}(B \setminus B_{n_0}))| \qquad (11)
\end{aligned}
$$

The first term of line (11) can be estimated as

$$
\begin{aligned}
& \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_{n_0}) - \mu(S^{-i}A)\nu(S^{-i}B)| \\
\le \; & \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_{n_0}) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_{n_0})| + \sup_{i\in\mathbb{N}} |\mu(S^{-i}A)||\mu(S^{-k}S^{-i}B_{n_0}) - \nu(B_{n_0})| \\
+ \; & \sup_{i\in\mathbb{N}} |\mu(S^{-i}A)||\nu(B_{n_0}) - \nu(B)|
\end{aligned}
$$

There exists a $k_1 \in \mathbb{N}$ such that $|\mu(S^{-k}S^{-i}B_{n_0}) - \nu(B_{n_0})| < \epsilon$ for all $k \ge k_1$, and there exists a $k_2 \in \mathbb{N}$ such that $\sup\limits_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B_{n_0}) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B_{n_0})| < \epsilon$ for all $k \ge k_2$. Thus, for all $k \ge \max\{k_1, k_2\}$, the first term of line (11) is small.

The second term of line (11) can be estimated as

$$
\begin{aligned}
& \sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}(B \setminus B_{n_0}))| \le \sup_{i\in\mathbb{N}} |\mu(S^{-k}S^{-i}(B \setminus B_{n_0}))| \\
\le \; & \sup_{i\in\mathbb{N}} |\mu(S^{-k}S^{-i}B) - \nu(B)| + |\nu(B) - \nu(B_{n_0})| + \sup_{i\in\mathbb{N}} |\nu(B_{n_0}) - \mu(S^{-k}S^{-i}B_{n_0})|
\end{aligned}
$$

which is small for all $k \ge \max\{k_0, k_1\}$. Hence, we have that

$$
\sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| < 6\epsilon \text{ for all } k \ge \max\{k_0, k_1, k_2\}.
$$

Thus $B \in \Lambda_A$. By the Dynkin $\pi - \lambda$ Theorem, for all $A \in \Pi$ and all $B \in \Sigma$ we have $\lim\limits_{k\to\infty} \sup\limits_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| = 0$. The same argument can be turned around to show that for a fixed $B \in \Sigma$ (9) holds for all $A \in \Sigma$. $\square$

Using Lemma 2.2, we obtain the following corollary of Theorem 2.1.

Corollary 2.3. *Let E be a separable metric space, μ be a probability measure on $\mathcal{B}(E)$, and $S : E \to E$ be Borel measurable. Let $\Pi \subset \Sigma$ be a π-system such that $\sigma(\Pi) = \Sigma$. Assume that*

1. *For every* $A, B \in \Pi$,
$$\sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}B) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}B)| \xrightarrow[k\to\infty]{} 0,$$
and
2. $(E, \mathcal{B}(E), \mu, S)$ *is asymptotically stationary with stationary limit* ν.

For every $n \in \mathbb{N}$ *define* $\mu_n : \mathcal{B}(E^n) \to [0,1]$ *by*
$$\mu_n(A) = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu\{x \in E : (S^{\sigma(1)}(x), ..., S^{\sigma(n)}(x)) \in A\}.$$
Then $(\mu_n)_{n\in\mathbb{N}}$ *is* ν*-chaotic.*

The next two remarks give sufficient conditions for the assumptions of Theorem 2.1 to be met.

Remark 2.4. *Let* (Ω, Σ, μ, S) *be a dynamical system which is mixing and* S *is* μ*-measure preserving. Then the assumptions 1 and 2 of Theorem 2.1 are satisfied for this dynamical system.*

Indeed, for every $A \in \Sigma$ we have

$$\begin{aligned}\sup_{i\in\mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}A) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}A)| &= \sup_{i\in\mathbb{N}} |\mu(S^{-i}(A \cap S^{-k}A)) - \mu(A)\mu(S^{-k}A)| \\ &= |\mu(A \cap S^{-k}A) - (\mu(A))^2| \xrightarrow[k\to\infty]{} 0\end{aligned}$$

where the second equality is valid because S is μ-measure preserving and the limit is valid because the dynamical system is mixing. Thus assumption 1 of Theorem 2.1 is satisfied. Also, (Ω, Σ, μ, S) is asymptotically stationary with stationary limit μ since S is μ-measure preserving. Thus assumption 2 of Theorem 2.1 is satisfied as well.

Remark 2.5. *Let* (Ω, Σ, μ, S) *be a dynamical system and let* $\Pi \subset \Sigma$ *be a* π*-system such that* $\sigma(\Pi) = \Sigma$.

(i) *If* $\lim_{k\to\infty} |\mu(S^{-k}A \cap B) - \mu(S^{-k}A)\mu(B)| = 0$ *holds for every* $A, B \in \Pi$, *then it holds for every* $A, B \in \Sigma$.
(ii) *If* $\mu(S^{-1}(A)) = \mu(A)$ *holds for all* $A \in \Pi$ *then it holds for all* $A \in \Sigma$.

See Shalizi and Kontorovich [15, Theorem 384] for the proof of part (i). The proof of part (ii) is very easy using Dynkin's $\pi - \lambda$ theorem.

We now present three examples of dynamical systems that satisfy the assumptions of Remark 2.4. The first example is called the "baker's map". The measure space for the baker's map is $([0,1]^2 := [0,1] \times [0,1], \mathcal{B}([0,1]^2), \mu)$ where μ is the Lebesgue measure restricted to $\mathcal{B}([0,1]^2)$. The map $S : [0,1]^2 \to [0,1]^2$ of the dynamical system is defined by

$$S(x,y) = \begin{cases} (2x, \frac{1}{2}y) & 0 \leq x < \frac{1}{2}, 0 \leq y \leq 1 \\ (2x-1, \frac{1}{2}y + \frac{1}{2}) & \frac{1}{2} \leq x \leq 1, 0 \leq y \leq 1 \end{cases}.$$

Lasota and Mackey [16, Example 4.3.1] prove that the baker's map is mixing by verifying Definition 1.1(ii) for all rectangles A, B with sides parallel to x and y axes. These rectangles form a π-system that generates the σ-algebra $\mathcal{B}([0,1] \times [0,1])$. Thus by Remark 2.5(i) the baker's map is mixing. Using the same π-system and Remark 2.5(ii) it is easy to verify that the baker's map is measure preserving.

Another example of a dynamical system which satisfies the assumptions of Remark 2.4 is the Anosov map, (also called the cat map). The measure space for the Anosov map is $([0,1)^2 := [0,1) \times [0,1), \mathcal{B}([0,1)^2), \mu)$ where μ is the Lebesgue measure restricted to $\mathcal{B}([0,1)^2)$. The map $S : [0,1)^2 \to [0,1)^2$ for the Anosov map is defined by

$$S(x,y) = (x+y, x+2y) \,(\text{mod } 1).$$

Lasota and Mackey prove that the cat map is mixing by using the Fibonacci sequence and Fourier transforms [16, Example 4.4.3]. Arnold and Avez show that the Anosov Map is measure preserving [1, Example 1.16].

We now present a construction of an infinite product of probability measures satisfying the assumptions of Remark 2.4. If $(\Omega_n, \Sigma_n, \mu_n)_{n\in\mathbb{N}}$ is a sequence of probability spaces then the Cartesian product $\prod_{n=1}^{\infty} \Omega_n$ can be naturally equipped with an infinite product of these measures, as defined by Kakutani [14]. Denote this infinite product probability space by $(\prod_{n=1}^{\infty} \Omega_n, \prod_{n=1}^{\infty} \Sigma_n, \prod_{n=1}^{\infty} \mu_n)$. The σ-algebra $\prod_{n=1}^{\infty} \Sigma_n$ is generated by the π-system $\prod_{n=1}^{<\infty} \Sigma_n$ consisting of all sets of the form $\prod_{n=1}^{\infty} A_n$ where $A_n \in \Sigma_n$ for all $n \in \mathbb{N}$ and $A_n = \Omega_n$ for all but finitely many n's. If $\Sigma_n = \Sigma$ for all $n \in \mathbb{N}$ then let $\Sigma^{<\infty}$ denote the π-system $\prod_{n=1}^{<\infty} \Sigma$. If $A = \prod_{n=1}^{\infty} A_n$ is such a set then we define $(\prod_{n=1}^{\infty} \mu_n)(A) = \prod_{n=1}^{\infty}(\mu_n(A_n))$. If $\Omega_n = \Omega$ and $\Sigma_n = \Sigma$ for all $n \in \mathbb{N}$ then $\prod_{n=1}^{\infty} \Omega_n$ is denoted by $\Omega^{\mathbb{N}}$, and $\prod_{n=1}^{\infty} \Sigma_n$ is denoted by $\Sigma^{\mathbb{N}}$. If moreover $\mu_n = \mu$ for all $n \in \mathbb{N}$ then $\prod_{n=1}^{\infty} \mu_n$ is denoted by $\mu^{\mathbb{N}}$. Assume that for every $n \in \mathbb{N}$ $(\Omega, \Sigma, \mu_n, S)$ is a dynamical system (i.e. in general we may allow different measures to be considered on the same σ-algebra Σ). Define $S^{\mathbb{N}} : \Omega^{\mathbb{N}} \to \Omega^{\mathbb{N}}$ by

$$S^{\mathbb{N}}((\omega_n)_{n\in\mathbb{N}}) = (S(\omega_{n+1}))_{n\in\mathbb{N}}.$$

Then $S^{\mathbb{N}}$ is measurable i.e. $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \prod_{n=1}^{\infty} \mu_n, S^{\mathbb{N}})$ is a dynamical system. Indeed, it is enough and easy to check that $(S^{\mathbb{N}})^{-1}(A) \in \Sigma^{\mathbb{N}}$ for every set A in the π-system $\Sigma^{<\infty}$. Obviously, if (Ω, Σ, μ, S) is a dynamical system and S is μ-measure preserving, then $S^{\mathbb{N}}$ is $\mu^{\mathbb{N}}$-measure preserving, (it is enough to be verified on sets of the π-system $\prod_{n=1}^{<\infty} \Sigma_n$, which is an easy task). Thus, if (Ω, Σ, μ, S) denotes either the baker's dynamical system, or the Anosov dynamical system defined above, then $S^{\mathbb{N}}$ is μ^N- measure preserving. We claim that for any dynamical system (Ω, Σ, μ, S), $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \mu^{\mathbb{N}}, S^{\mathbb{N}})$ is always mixing. By Remark 2.5(i) this claim is enough and easy to be verified for sets A, B in the π-system $\Sigma^{<\infty}$. Indeed if $A = \prod_{n=1}^{\infty} A_n$, $B = \prod_{n=1}^{\infty} B_n$ where $A_n, B_n \in \Sigma$ for all n and $A_n = B_n = \Omega$ for all $n > m$, then

$$\mu^{\mathbb{N}}((S^{\mathbb{N}})^{-k} A) = \prod_{n=1}^{m} \mu(S^{-k} A_n), \quad \mu^{\mathbb{N}}(B) = \prod_{i=1}^{m} \mu(B_i),$$

and

$$\mu^{\mathbb{N}}((S^{\mathbb{N}})^{-k} A \cap B) = \prod_{i=1}^{m} \mu(B_i) \prod_{n=1}^{m} \mu(S^{-k} A_n) \quad \text{for all } k > m.$$

Hence

$$|\mu^{\mathbb{N}}((S^{\mathbb{N}})^{-k} A \cap B) - \mu^{\mathbb{N}}((S^{\mathbb{N}})^{-k} A)\mu^{\mathbb{N}}(B)| = 0 \quad \text{for all } k > m.$$

Thus if (Ω, Σ, μ, S) is a dynamical system such that S is μ-measure preserving, then we obtain that $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \mu^{\mathbb{N}}, S^{\mathbb{N}})$ is a dynamical system that satisfies the assumptions of Remark 2.4.

In all the examples that we have mentioned so far, the map of the dynamical system is measure preserving. Such maps are trivially asymptotically stationary with stationary limits being equal to the original measure. We now describe how infinite product probability measures can be used to give examples of dynamical systems that satisfy the assumptions of Theorem 2.1 without the map of the dynamical system being measure preserving. In this example the stationary limit of the dynamical system is different than the original measure. Let $\Omega = [0, 1]$ and $\Sigma = \mathcal{B}([0, 1])$. For every $k \in \mathbb{N}$ define a density function $\phi_k : [0, 1] \to \{0, 1, 2\}$ by

$$\phi_k(x) = \chi_{[0,1-\frac{1}{2^{k-1}}]}(x) + 2\chi_{(1-\frac{1}{2^{k-1}},1-\frac{1}{2^k}]}(x)$$

and the probability measure $\mu_k : \Sigma \to [0, 1]$ by

$$\mu_k(A) := \int_A \phi_k(x)dx \quad \text{for all } A \in \Sigma.$$

Consider the probability space $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \prod_{k=1}^{\infty} \mu_k)$ and in order to make the notation easier let $\mathbb{M} = \prod_{k=1}^{\infty} \mu_k$ and $\mathbb{L} = \lambda^{\mathbb{N}}$ where λ is the Lebesgue measure on $[0, 1]$. Consider a map $S : \Omega \to \Omega$ such that $S^{\mathbb{N}}$ is $\mathbb{L}$-measure preserving but not $\mathbb{M}$-measure preserving (this is valid for example when S is the identity map). In order to make the notation easier let $\mathbb{S} = S^{\mathbb{N}}$. We claim that the dynamical system $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \mathbb{M}, \mathbb{S})$ satisfies the assumptions of Theorem 2.1. Indeed, in order to verify assumption 2 of Theorem 2.1, we prove that $\mathbb{M}$ is asymptotically stationary with stationary limit equal to $\mathbb{L}$.

Fix any $A \in \Sigma^{\mathbb{N}}$. Define for each $k \in \mathbb{N}$ the set $C_k := [0, 1]^k \times [0, 1 - \frac{1}{2^k}] \times [0, 1 - \frac{1}{2^{k+1}}] \times [0, 1 - \frac{1}{2^{k+2}}] \times \cdots$. Since $C_k = \bigcap_{n=0}^{\infty} \big([0, 1]^k \times [0, 1 - \frac{1}{2^k}] \times \cdots \times [0, 1 - \frac{1}{2^{k+n}}] \times [0, 1]^{\mathbb{N}}\big)$, we have that $C_k \in \Sigma^{\mathbb{N}}$, and if $B \in \Sigma^{\mathbb{N}}$ with $B \subset C_k$ then $\mathbb{M}(B) = \mathbb{L}(B)$. Thus for any $A \in \Sigma^{\mathbb{N}}$ we have that

$$\begin{aligned} & |\mathbb{M}(\mathbb{S}^{-k}A) - \mathbb{L}(A)| = |\mathbb{M}(\mathbb{S}^{-k}A) - \mathbb{L}(\mathbb{S}^{-k}B)| \\ = \; & |\mathbb{M}((\mathbb{S}^{-k}A) \cap C_k) + \mathbb{M}((\mathbb{S}^{-k}A) \setminus C_k) - \mathbb{L}((\mathbb{S}^{-k}A) \cap C_k) - \mathbb{L}((\mathbb{S}^{-k}A) \setminus C_k)| \\ = \; & |\mathbb{M}((\mathbb{S}^{-k}A) \setminus C_k) - \mathbb{L}((\mathbb{S}^{-k}A) \setminus C_k)| \leq 2\left(1 - \prod_{s=k}^{\infty}(1 - \frac{1}{2^s})\right) \xrightarrow[k\to\infty]{} 0, \end{aligned}$$

where the last inequality is valid because $\mathbb{M}(C_k) = \mathbb{L}(C_k) = \prod_{s=k}^{\infty}(1 - \frac{1}{2^s})$, hence $\mathbb{M}([0,1]^{\mathbb{N}} \setminus C_k) = \mathbb{L}([0,1]^{\mathbb{N}} \setminus C_k) = 1 - \prod_{s=k}^{\infty}(1 - \frac{1}{2^s})$. This verifies assumption 2 of Theorem 2.1.

Now, in order to verify assumption 1 of Theorem 2.1 we use Lemma 2.2. Fix sets $A, B \in \prod_{n=1}^{<\infty} \mathcal{B}([0,1])$. Then $A = \prod_{n=1}^{\infty} A_n$ where $A_n \in \mathcal{B}([0,1])$ and there exists $N \in \mathbb{N}$ such that $A_n = [0,1]$ for all $n > N$. Then by the definition of the infinite product measure we have that for all $k > N$

$$\mathbb{M}((\mathbb{S}^{-i}A) \cap (\mathbb{S}^{-k}\mathbb{S}^{-i}B)) = \mathbb{M}(\mathbb{S}^{-i}A)\mathbb{M}(\mathbb{S}^{-k}\mathbb{S}^{-i}B) \text{ for all } i \in \mathbb{N}.$$

Hence (9) is valid, and the assumptions of Lemma 2.2 are met. This means assumption 1 of Theorem 2.1 is valid.

3. Proof of the Main Result

Many times when deciding whether a sequence of measures is ν-chaotic, it is easier to show one of the equivalent formulations of chaos. Sznitman proves various equivalences to the definition of chaos which we list below.

Theorem 3.1. [23, Proposition 2.2] *Let E be a separable metric space, $(\mu_n)_{n=1}^{\infty}$ a sequence of symmetric probability measures on E^n (as in Definition 1.2), and ν be a probability measure on E. The following are equivalent:*

1. *The sequence $(\mu_n)_{n=1}^{\infty}$ is ν-chaotic (as in Definition 1.3).*
2. *The function $X_n : E^n \to M(E)$ defined by $X_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ (where δ_x stands for the Dirac measure at x) converges in law with respect to μ_n to the constant random variable ν, i.e. for every $g \in C_b(E)$ we have that*

$$\int_{E^n} |(X_n - \nu)g|^2 d\mu_n \xrightarrow[n\to\infty]{} 0,$$

where $C_b(E)$ stands for the space of bounded continuous scalar valued functions on E.

3. *The sequence $(\mu_n)_{n=1}^{\infty}$ satisfies Definition 1.3 with $k = 2$.*

In order to construct examples of sequences of symmetric probability measures satisfying condition 2 of Theorem 3.1 we will show that it is sufficient to construct a sequence of probability measures (not necessarily symmetric) which satisfy the same condition. Given a measurable space (E, Σ), $n \in \mathbb{N}$, and a probability measure μ_n on the product space (E^n, Σ^n), we define a symmetric probability measure μ_n^{sym} on (E^n, Σ^n) in the following way: For each $\sigma \in \Sigma_n$ define $\Pi_\sigma : E^n \to E^n$ by $\Pi_\sigma(x_1, ..., x_n) = (x_{\sigma(1)}, ..., x_{\sigma(n)})$, and define the probability measure $\mu_n^\sigma : \Sigma^n \to [0, 1]$ by $\mu_n^\sigma(A) = \mu_n(\Pi_\sigma(A))$ for each $A \in \Sigma^n$. It is easy to verify that for any bounded and measurable function $f : E^n \to \mathbb{C}$ we have that

$$\int_{E^n} f d\mu_n^\sigma = \int_{E^n} f^\sigma d\mu_n$$

where f^σ is defined as in (4). The symmetric probability measure $\mu_n^{\text{sym}} : \Sigma^n \to [0, 1]$ is then defined by

$$\mu_n^{\text{sym}}(A) = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \mu_n^\sigma(A) \text{ for each } A \in \Sigma^n.$$

For any bounded and measurable function $f : E^n \to \mathbb{C}$ which is symmetric (as in (3)),

$$\int_{E^n} f d\mu_n^{\text{sym}} = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \int_{E^n} f d\mu_n^\sigma = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \int_{E^n} f^\sigma d\mu_n = \int_{E^n} f d\mu_n. \tag{12}$$

For a fixed $g \in C_b(E)$ and $\nu \in M(E)$ if we apply (12) for $f := |(X_n - \nu)(g)|^2 : E^n \to \mathbb{C}$ (which is obviously bounded, measurable, and symmetric), we obtain the following.

Remark 3.2. *Let E be a separable metric space, μ_n be a probability measure on $\mathcal{B}(E^n)$, and ν be a probability measure on $\mathcal{B}(E)$. For any fixed $g \in C_b(E)$ we have that*

$$\int_{E^n} |(X_n - \nu)g|^2 d\mu_n \xrightarrow[n\to\infty]{} 0 \quad \textit{if and only if} \quad \int_{E^n} |(X_n - \nu)g|^2 d\mu_n^{sym} \xrightarrow[n\to\infty]{} 0.$$

In order to prove Theorem 2.1, we will also need the following.

Proposition 3.3. *Let E be a separable metric space, μ be a probability measure on $\mathcal{B}(E)$, and $S : E \to E$ be a Borel measurable map. Assume that*

1. *For every $A \in \mathcal{B}(E)$,*
$$\sup_{i \in \mathbb{N}} |\mu(S^{-i}A \cap S^{-k}S^{-i}A) - \mu(S^{-i}A)\mu(S^{-k}S^{-i}A)| \xrightarrow[k\to\infty]{} 0,$$
and
2. *$(E, \mathcal{B}(E), \mu, S)$ is asymptotically stationary with stationary limit ν.*

For every $n \in \mathbb{N}$ define $\mu_n : \mathcal{B}(E^n) \to [0,1]$ by

$$\mu_n(A) = \mu\{x \in E : (S(x), ..., S^n(x)) \in A\}.$$

Then for every $g \in C_b(E)$,

$$\int_{E^n} |(X_n - \nu)g|^2 d\mu_n \xrightarrow[n\to\infty]{} 0.$$

Notice that for the measure μ_n defined in Proposition 3.3, and for each $A \in \mathcal{B}(E^n)$,

$$\begin{aligned}
\mu_n^{\mathrm{sym}}(A) &= \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu_n^{\sigma}(A) = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu_n(\Pi_\sigma(A)) \\
&= \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu\{x \in E : (S(x), ..., S^n(x)) \in \Pi_\sigma(A)\} \\
&= \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu\{x \in E : (S^{\sigma^{-1}(1)}(x), ..., S^{\sigma^{-1}(n)}(x)) \in A\} \\
&= \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \mu\{x \in E : (S^{\sigma(1)}(x), ..., S^{\sigma(n)}(x)) \in A\}. \qquad (13)
\end{aligned}$$

Now the proof of Theorem 2.1 follows immediately from Proposition 3.3, Remark 3.2, the fact that μ_n^{sym} is a symmetric probability measure, Theorem 3.1, and (13). It remains to prove Proposition 3.3.

Proof of Proposition 3.3. First, let $g := \chi_{E_1}$ for some $E_1 \in \mathcal{B}(E)$. Then, using that for $1 \le i < j \le n$, we have

$$\int_{E^n} g(x_i)g(x_j)d\mu_n = \mu_n(E^{i-1} \times E_1 \times E^{j-i-1} \times E_1 \times E^{n-j}) = \mu(S^{-i}(E_1) \cap S^{-j}(E_1))$$

we can write

$$\begin{aligned}
\int_{E^n} |(X_n-\nu)(g)|^2\,d\mu_n &= \int_{E^n}\left|\frac{1}{n}\sum_{i=1}^n g(x_i)-\int_E g d\nu\right|^2 d\mu_n \\
&= \frac{1}{n^2}\sum_{1\le i,j\le n}\int_{E^n} g(x_i)g(x_j)d\mu_n - \frac{2\nu(E_1)}{n}\sum_{i=1}^n\int_{E^n} g(x_i)d\mu_n + (\nu(E_1))^2 \\
&= \frac{1}{n^2}\sum_{1\le i,j\le n}\mu(S^{-i}(E_1)\cap S^{-j}(E_1)) - \frac{2\nu(E_1)}{n}\sum_{1\le i\le n}\mu(S^{-i}(E_1)) + (\nu(E_1))^2 \\
&= \frac{2}{n^2}\sum_{1\le i<j\le n}\mu(S^{-i}(E_1)\cap S^{-(j-i)}(S^{-i}(E_1))) \qquad (14)\\
&+ \frac{1}{n^2}\sum_{1\le i\le n}\mu(S^{-i}(E_1)) - \frac{2\nu(E_1)}{n}\sum_{1\le i\le n}\mu(S^{-i}(E_1)) + (\nu(E_1))^2
\end{aligned}$$

We have

$$\frac{1}{n^2}\sum_{1\le i\le n}\mu(S^{-i}(E_1)) \le \frac{1}{n^2}\sum_{1\le i\le n} 1 \xrightarrow[n\to\infty]{} 0$$

and by assumption 2,

$$\frac{2\nu(E_1)}{n}\sum_{1\le i\le n}\mu(S^{-i}(E_1)) \xrightarrow[n\to\infty]{} 2(\nu(E_1))^2$$

Also, line (14) can be written as

$$\begin{aligned}
&\frac{2}{n^2}\sum_{1\le i<j\le n}\mu(S^{-i}(E_1)\cap S^{-(j-i)}(S^{-i}(E_1))) \\
&= \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\left[\mu\left(S^{-i}(E_1)\cap S^{-k}(S^{-i}(E_1))\right) - \mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))\right] \qquad (15)\\
&+ \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1))). \qquad (16)
\end{aligned}$$

First, let us focus on line (15). Let $\epsilon > 0$. By assumption 1 of Theorem 2.1, there exists $K_0 \in \mathbb{N}$ such that if $k \ge K_0$ then

$$|\mu(S^{-i}(E_1)\cap S^{-k}(S^{-i}(E_1))) - \mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))| < \epsilon$$

for every i. Thus for $n > K_0 + 1$ we have that line (15) is less than or equal to

$$\frac{2}{n^2}\sum_{k=1}^{K_0}\sum_{i=1}^{n-k}\left|\mu\left(S^{-i}(E_1)\cap S^{-k}(S^{-i}(E_1))\right)-\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))\right|+\frac{2}{n^2}\sum_{k=K_0+1}^{n-1}\sum_{i=1}^{n-k}\epsilon.$$

Since the first double sum has at most $K_0^2+\frac{K_0 n}{2}$ terms, the second double sum has at most $\frac{(n-K_0)n}{2}$ terms, and $0 \leq \left|\mu\left(S^{-i}(E_1)\cap S^{-k}(S^{-i}(E_1))\right)-\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))\right| \leq 2$, we have that line (15) is less than or equal to

$$\frac{4(K_0^2+K_0 n/2)}{n^2} \quad + \quad \frac{\epsilon(n-K_0)n}{n^2} \xrightarrow[n\to\infty]{} \epsilon.$$

Now we will focus on line (16). By assumption 2, there exists $N_0 \in \mathbb{N}$ such that if $n \geq N_0$ then

$$|\mu(S^{-n}(E_1) - \nu(E_1)| < \epsilon.$$

Hence, line (16) is equal to

$$\frac{2}{n^2}\sum_{i=1}^{N_0}\sum_{k=1}^{n-i}\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))+\frac{2}{n^2}\sum_{i=N_0+1}^{n-1}\sum_{k=1}^{n-i}\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1)))$$

The first double sum has at most $N_0^2 + \frac{N_0 n}{2}$ terms, and thus,

$$\frac{2}{n^2}\sum_{i=1}^{N_0}\sum_{k=1}^{n-i}\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1))) \leq \frac{2(N_0^2+N_0 n/2)}{n^2} \xrightarrow[n\to\infty]{} 0$$

The second double sum can be rewritten as

$$\begin{aligned}
& \frac{2}{n^2}\sum_{i=N_0+1}^{n-1}\sum_{k=1}^{n-i}\mu(S^{-i}(E_1))\mu(S^{-k}(S^{-i}(E_1))) \\
= \quad & \frac{2}{n^2}\frac{(n-1-N_0)^2}{2}(\nu(E_1))^2 \\
+ \quad & \frac{2}{n^2}\sum_{i=N_0+1}^{n-1}\sum_{k=1}^{n-i}\left[\mu(S^{-i}E_1)[\mu(S^{-k}S^{-i}E_1)-\nu(E_1)]+[\mu(S^{-i}E_1)-\nu(E_1)]\nu(E_1)\right].
\end{aligned}$$

We have

$$\frac{2}{n^2}\frac{(n-1-N_0)^2}{2}(\nu(E_1))^2 \xrightarrow[n\to\infty]{} (\nu(E_1))^2$$

and

$$\begin{aligned}
&\frac{2}{n^2}\sum_{i=N_0+1}^{n-1}\sum_{k=1}^{n-i}\left[\mu(S^{-i}E_1)[\mu(S^{-k}S^{-i}E_1)-\nu(E_1)]+[\mu(S^{-i}E_1)-\nu(E_1)]\nu(E_1)\right] \\
\leq\ &\frac{2}{n^2}\frac{(n-1)^2}{2}\epsilon+\frac{2}{n^2}\frac{(n-1)^2}{2}\epsilon \xrightarrow[n\to\infty]{} (\nu(E_1))^2+2\epsilon.
\end{aligned}$$

Hence line (16) converges to $(\nu(E_1))^2$ as $n\to\infty$. This shows

$$\int_{E^n}|(X_n-\nu)(g)|^2\,d\mu_n \xrightarrow[n\to\infty]{} 0$$

for $g=\chi_{E_1}$.

Now, consider the simple function $g := \sum_{k=1}^{K}\alpha_k\chi_{E_k}$. We have

$$\begin{aligned}
\left(\int_{E^n}|(X_n-\nu)\,(g)|^2\,d\mu_n\right)^{1/2} &= \left(\int_{E^n}\left|(X_n-\nu)\left(\sum_{k=1}^{K}\alpha_k\chi_{E_k}\right)\right|^2 d\mu_n\right)^{1/2} \\
= \left(\int_{E^n}\left|\sum_{k=1}^{K}\alpha_k\left[(X_n-\nu)\chi_{E_k}\right]\right|^2 d\mu_n\right)^{1/2} &\leq \sum_{k=1}^{K}|\alpha_k|\left(\int_{E^n}|(X_n-\nu)\chi_{E_k}|^2\,d\mu_n\right)^{1/2} \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

since the limit is zero for each characteristic function and we have a finite sum.

Finally, let $g\in C_b(E)$ and $\epsilon>0$. There exists a simple function G such that $|g(x)-G(x)|<\epsilon$ for all $x\in E$. We have that

$$
\begin{aligned}
\left(\int_{E^n}|(X_n-\nu)(g-G)|^2\,d\mu_n\right)^{1/2} &\le \left(\int_{E^n}|X_n(g-G)|^2 d\mu_n\right)^{1/2}+\left(\int_{E^n}|\nu(g-G)|^2 d\mu_n\right)^{1/2}\\
&= \left(\int_{E^n}\left|\frac{1}{n}\sum_{i=1}^{n}(g-G)(x_i)\right|^2 d\mu_n\right)^{1/2}+|\nu(g-G)|\\
&\le \frac{1}{n}\sum_{i=1}^{n}\left(\int_{E^n}|(g-G)(x_i)|^2 d\mu_n\right)^{1/2}+\left|\int_E (g-G)d\nu\right|\\
&\le \sum_{i=1}^{n}\frac{\epsilon}{n}+\epsilon=2\epsilon
\end{aligned}
$$

and since G is a simple function,

$$\left(\int_{E^n}|(X_n-\nu)G|^2 d\mu_n\right)^{1/2}\xrightarrow[n\to\infty]{}0.$$

Therefore by the triangle inequality on $L^2(E^n,d\mu_n)$ norm we obtain since ϵ is arbitrary,

$$\left(\int_{E^n}|(X_n-\nu)g|^2 d\mu_n\right)^{1/2}\xrightarrow[n\to\infty]{}0.$$

□

References

[1] V.I. Arnold, A. Avez, *Ergodic properties of classical mechanics*, Benjamin (1968).

[2] V. Bogachev, *Measure Theory* Volume 1, Springer Verlag, (2007).

[3] E.A. Carlen, M.C. Carvalho, J. Le Roux, M. Loss, C. Villani, *Entropy and chaos in the Kac model*, Kinet. Relat. Models **3** (1) (2010), 85-122.

[4] E. Carlen, M.C. Carvalho, M. Loss, *Kinetic theory and the Kac master equation*, Entropy & Quantum II, Contemp. Math. **552**, (2011), 1-20.

[5] K. Carrapatoso, Quantitative and qualitative Kac's chaos on the Boltzmann's sphere, Ann. Inst. H. Poincar Probab. Statist. **51**, Number 3 (2015), 993-1039.

[6] G. Casati, B. Chirikov (Editors), *Quantum Chaos, between order and disorder*, Cambridge University Press, (1995).

[7] M. Giannoni, A. Voros, J. Zinn-Justin (Editors), *Chaos and Quantum Physics (Les Houches)*, North Holland, (1991).

[8] G. Graham, S. Méléard, *Stochastic particle approximations for generalized Boltzmann models and convergence estimates*, Ann. Propab. **25** (1997), 115-132.

[9] F.A. Grünbaum, *Propagation of chaos for the Boltzmann equation*, Arch. Rational Mech. Anal. **42** (1971), 323-345.

[10] M. Hauray, S. Mischler, On Kac's chaos and related problems, J. Funct. Anal. **266** (2014), 6055-6157.

[11] W. Dieter Heiss (Ed.) *Chaos and Quantum Chaos*, Proceedings of the Eighth Chris Engelbrecht Summer School in Theoretical Physics. Held at Blydepoort, Eastern Transvaal South Africa, 13-24 January 1992, Lecture Notes in Physics, Springer Verlag.

[12] D.P. Jonhson, *On a class of stochastic processes and its relationship to infinite particle gases*, Trans. Amer. Math. Soc. **132** (1968), 275-295.

[13] M. Kac, *Foundations of Kinetic Theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability: held at the Statistical Laboratory, University of California, (1956).

[14] S. Kakutani, *On Equivalence of Infinite Product Measures*, Annals of Mathematics, Vol. **49**, No. 1, (Jan. 1948), p.214-224.

[15] A. Kontorovich, C. Shalizi, *Almost None of the Theory of Stochastic Processes*, Version 0.1.1, July 3, 2010, available in author's website: https://www.stat.cmu.edu/ cshalizi/almost-none/

[16] A. Lasota, M.C. Mackey, *Probabilistic Properties of Deterministic Systems*, Cambridge University Press, (1985).

[17] H.P. McKean, *Propagation of chaos for a class of non-linear parabolic equations*, Stochastic Differential Equations, (Lecture Series in Differential Equations Session 7, Catholic University, 1967), 41-57, Air Force Office Sci. Res. Arlington, Va.

[18] S. Mischler, *Le programme de Kac sur les limites de champ moyen*, in: Séminaire EDP-X, Décembre 2010.

[19] S. Mischler, C. Mouhot, *Kac's program in kinetic theory*, Invent. Math. **193** (1) (2014) 1-147.

[20] H. Murata, *Propagation of chaos for Boltzmann like equations of non-cut off type in the plane*, Hiroshima Math. J. **7** (1977), 479-515.

[21] Y. Sinai, *Ergodic Teory*, Acta Physica Austriaca, Suppl. X, (1973), pp. 575-608.

[22] A-S. Sznitman, *Équations de type de Boltzmann spatialement homogènes*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **66** (1984), 559-592.

[23] A-S. Sznitman, *Topics in Propagation of Chaos*, Lecture Notes in Mathematics (2006), pp. 165-251.

[24] H. Tanaka, *Propagation of chaos for certain Markov processes of jump type with non-linear generators*, Part I, Part II, Proc. Japan Acad. **45** (1969), 449-452, 598-599.

[25] T. Ueno, *A class of Markov processes with non-linear bounded generators*, Japan J. Math. **38**, (1969), 19-38.

Department of Mathematics, University of South Carolina, Columbia, SC 29208
E-mail address: giorgis@math.sc.edu, musulin@math.sc.edu